

AFRL-IF-RS-TR-2001-26
Final Technical Report
March 2001



TECHNIQUES FOR ASSOCIATIVE RETRIEVAL

Stanford University

Xiaochun Li and Shuaib Uddin Arshad

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

20010508 015

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2001-26 has been reviewed and is approved for publication.

APPROVED:



PETER J. COSTIANES
Project Engineer

FOR THE DIRECTOR:



JOSEPH CAMERA, Chief
Information & Intelligence Exploitation Division
Information Directorate

If your address has changed or if you wish to be removed from the Air Force Research Laboratory Rome Research Site mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/IFED, 32 Brooks Road, Rome, NY 13441-4114. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE MARCH 2001		3. REPORT TYPE AND DATES COVERED Final Apr 00 - Oct 00
4. TITLE AND SUBTITLE TECHNIQUES FOR ASSOCIATIVE RETRIEVAL			5. FUNDING NUMBERS C - F30602-00-1-0522 PE - 63726F PR - 281E TA - 00 WU - P1	
6. AUTHOR(S) Xiaochun Li and Shuaib Uddin Arshad				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Stanford University Department of Electrical Engineering 250 Durand Building Stanford CA 94305-4035			8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/IFED 32 Brooks Road Rome NY 13441-4114			10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-IF-RS-TR-2001-26	
11. SUPPLEMENTARY NOTES Air Force Research Laboratory Project Engineer: Peter J. Costianes/IFED/(315) 330-4030				
12a. DISTRIBUTION AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This report is a comparison and documentation of techniques of associative retrieval. Techniques presented include conventional database, artificial neural networks, feature-based digital image associative search, holographic associative memory, and content-addressable processors.				
14. SUBJECT TERMS Associative Retrieval, Holographic Associative Memory, Neural Networks, Patter Recognition			15. NUMBER OF PAGES 36	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

Comparison and Documentation of Techniques of Associative Retrieval

Contents

1. Overview of five associative retrieval methods	1
1.1 Classification of information in associative computing—three kinds of information	1
1.2 Comparison of five associative methods	1
1.2.1 Conventional database	1
1.2.2 Artificial neural network (ANN)	1
1.2.3 Feature-based digital image associative search	2
1.2.4 Holographic associative memory (HAM)	2
1.2.5 Content-addressable processor	3
2. Conventional database	3
2.1 References for conventional database	4
3. Neural networks for nonlinear pattern recognition	5
3.1 Nonlinear pattern recognition	5
3.2 Artificial neural network, learning and parallel processing	5
3.3 Neural networks for nonlinear processing	6
3.4 Holographic optical neural network system	8
3.4.1 Massive interconnection requirement	8
3.4.2 Implementation of a holographic optical neural network	9
3.4.3 Pattern recognition	11
3.5 Comments	11
3.6 References for neural networks for nonlinear pattern recognition	11
4. Digital image associative search	13
4.1 Retrieval of image information	13
4.2 Content based image retrieval	13
4.3 Approaches	14
4.4 Query classes	14
4.5 Using color to retrieve image	15
4.6 Color histogram	15
4.7 Advantages of color histogram	16
4.8 Limitation of color histogram	16
4.9 Color coherence vectors (CCV)	17
4.10 CCV computation	17
4.11 Comparing CCV's	18
4.12 Computational efficiency	18
4.13 References for digital image associative search	19
4.14 Web Sites for Content-Based Image Retrieval	20

5. Volume holographic associative memory and other optical correlators	24
5.1 Volume holographic associative memory	24
5.2 References for volume holographic associative memory	24
5.3 Other optical correlators	24
5.4 References for other optical correlators	25
6. Content-addressable processor	27
6.1 Totally digital and electronic processor	27
6.2 Optical processor	27
6.3 References for content-addressable processor	27

List of Figures

Figure No.		Page
2.1	Database in a Client/Server system	4
2.2	Database in a web-based system	4
3.1	Object A, B and C in a two-feature space. (a) A, B and C are linearly separable by two hyperplanes, D1 and D2. (b) A, B and C are not linearly separable.	5
3.2	Operational flow of a multi-layer neural network. It consists of a series of matrix-vector multiplication and nonlinear thresholding.	6
3.3	Comparison of a linear classifier, a two-layer NN and a three-layer NN. The thin lines represent the decision plane boundary.	8
3.4	Recording and reconstruction of a large scale holographic optical neural network.	10

Comparison and Documentation of Techniques of Associative Retrieval¹

Xiaochun Li and Shuaib Uddin Arshad

Stanford University

1. Overview of five associative retrieval methods

In this section we identify three types of information to be processed by associative retrieval methods, then identify five associative retrieval methods. The advantages, disadvantages and applications of each method will be addressed in subsequent sections of the report.

1.1 Classification of information in associative computing—three kinds of information

The information to be processed in associative computing can be divided into three types, plain text, graphical images and natural images.

- (1) **Plain text.** Examples include scientific and business documentation, information libraries (employee, product, student...), etc
- (2) **Graphical image.** Examples include maps, handwritten characters, circuit drawings, etc.
- (3) **Natural image.** Examples include medical diagnostic image, infrared image, fingerprint, satellite landscape image, etc.

1.2 Comparison of five associative methods

Associative retrieval methods can be divided into five categories as follows: conventional database, artificial neural network, feature-based digital image associative search, holographic associative memory and content-addressable processor. These are described below.

1.2.1 Conventional database

The theory and technology of conventional database has been successfully developed in the past 30 years and it has become a mature data retrieval technique. Many commercial products such as Oracle 8, Microsoft SQL Server are available. All conventional database have some form of strict data structure, and SQL (structured query language) is used to maintain this structure. For this reason conventional database is also called relational database. Since the operation of conventional database is based on software and keywords, it is well applicable to the management of structured textual information, but it has difficulty handling image information directly. Examples of the application of conventional database include library information search systems, and internet-based search engines such as Yahoo, Google, Excite, etc.

1.2.2 Artificial neural network (ANN)

The associative computing principle of ANN is based on 1) the emulation of biological neural systems and 2) Kolmogorov's theorem. In Kolmogorov's theorem it has

¹ Report prepared under Contract No. F30602-00-1-0522 with Air Force Research Laboratory

been proved that any continuous multivariate function can be expressed by a finite number of single variable continuous functions and a set of parameters. This implies that a finite architecture will be able to imitate the input-output behavior of any complex system. Because the same set of component functions can be used to construct any function, the function can be represented just by adjusting a set of variable parameters. The so-called learning ability of ANN is in fact the process of finding the appropriate set of parameters. ANN has many advantages in information processing. Patterns of arbitrary complexity can be learned through general purpose learning algorithms. Once learned, massive numbers of patterns can be rapidly searched and retrieved. It is also robust and highly parallel.

Though it has been predicted since its inception that ANN will play an important role in associative computing, the success of ANN over the past 50 years are primarily limited to adaptive information classification or filtering and fast approximate optimization (like the Hopfield model). For an ANN based associative memory, the input query object has to be statistically dominant, usually more than 60% of the stored (learned) pattern. As we know, however, for a practical search operation, the input search key size may be very small, as, for example, finding a small target in cluttered background.

1.2.3 Feature-based digital image associative search

This technique is in fact the functional extension of conventional database so that conventional database can process image information. It is based on annotated symbolic models of image contents, and uses the textual representations in terms of a set of pre-decided attributes (such as color, shape, size, etc.) and their values. Machine detectable features such as geometric moments, triangular cover, points of maximum curvature are often used to help in automating the model extraction process. These symbolic descriptions are then stored into some form of data structure. Once the data structure is available, various conventional database search techniques are used to search it. As can be seen, the search operation is performed on the image-feature vectors, not on the image itself. This retrieval method is also called pseudo-content addressable memory.

Model-based content addressable memory lends itself readily to images of graphical nature, where the image concept is well designed and extensive domain specific modeling and human involvement is possible. For natural images, the model extraction (translation of the "meaning" of an image or its objects into a structured model) itself is a formidable task, because of the inherent amorphousness of image information. Examples of the application of this technique include IBM's QBIC (Query by image content) database and PACs used in hospitals.

1.2.4 Holographic associative memory

In holographic associative memory (HAM) the association ability is an inherent property of holographic data storage, arising from the fact that the reference beam used to record one of a number of holograms may be reconstructed by using a small portion of the object beam as a search key. The high parallelism of HAM makes it possible to achieve high data rates for storage and retrieval. Due to the strict Bragg selectivity of volume holograms, this data retrieval method works best for objects having a fixed position within a holographic record. Holographic data storage has not found practical

application mainly because of the lack of appropriate recording materials having large storage capacity, long storage time, low scattering, high dynamic range and high resolution. However there has been significant progress in the past few years, and new media and optoelectronic devices are available that may allow commercial development. HAM is not as flexible as computer-based conventional database, but it may provide a useful data management tool in a holographic data storage system.

1.2.5 Content-addressable processor

Unlike the associative methods discussed above, content-addressable processors perform only content comparison or other logic operations on the input data. They do not store any information. There are two types of content-addressable processor.

1.2.5.1 Totally digital and electronic processor

These processors are in fact single semiconductor chips. They have been used as address filters or address translators in ATM and Ethernet based systems. NetLogic Microsystems Inc, Mountain View, CA, for example, is one provider of commercial products. Using a NetLogic chip for a 16K×64 routing table in a 12 port Ethernet system, the address filtering, and source and destination search time can be less than 30ns. Usually the capacity of this type of content-addressable processor is very small. Detailed information on NetLogic's product information can be found at <http://www.netlogicmicro.com/>.

1.2.5.2 Optical processor

A good example of an optical content-addressable processor is the one recently proposed by A. Louri et al at ECE Dept, the University of Arizona, which is called multiwavelength optical content-addressable parallel processor (MW-OCAPP). The MW-OCAPP is designed to provide efficient parallel data retrieval and processing by means of moving the bulk of database operations from electronics to optics. Polarization and wavelength-encoding have been proposed to enhance processing parallelism. 11 relational database primitive operations can be realized in MW-OCAPP. The main problems of these optical processors seem to be the difficulty in manufacturing high speed SLMs, the low data transfer rate from computer to SLM, the computation inaccuracy, and bulk system size.

2. Conventional database^[1-3]

The theory and technology of conventional database has been successfully developed in the past 30 years, and it has become a mature data retrieval technique. Many commercial products such as Oracle 8i, Microsoft SQL Server and Informix Dynamic Server2000 are available in market.

A database has a well defined structure. Not only does it contain the data of interest, it also contains metadata that describes the structure of the data within a database. Because the database contains a description of its own structure, it is said to be self-describing. Relational databases have become popular in the past 30 years because of their flexible structure. Relational databases are dynamic and size scalable. When the structure of a relational database is changed, the application program code does not need any modification. In addition to the updatable structure, the source data can also be

dynamically updated and deleted. The industry standard computer language to maintain relational database structure is SQL (Structured Query Language). The syntax of SQL is much like standard English, and easily understood. The operation of conventional database is through software manipulation, and can be based on arbitrarily logical and arithmetic combinations. A conventional database is constructed on and operated through either a client/server system or a web-based system. As shown in Fig.2.1 and Fig.2.2, included in the systems are a large capacity data storage device (for source data), high performance computer (for server), application program (DBMS, browser, client and server extensions) and communication channel. All this hardware and software are becoming increasingly mature and thus cheaper.

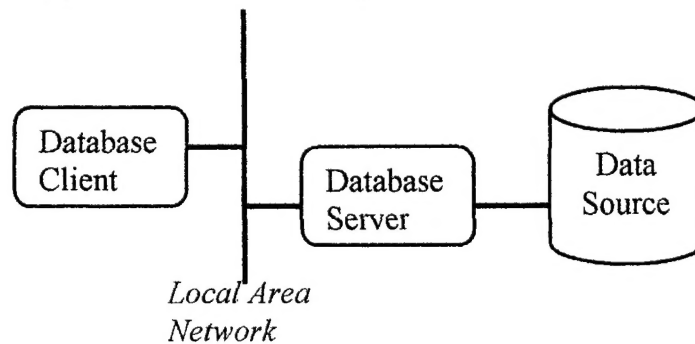


Fig.2.1 Database in a Client/Server system

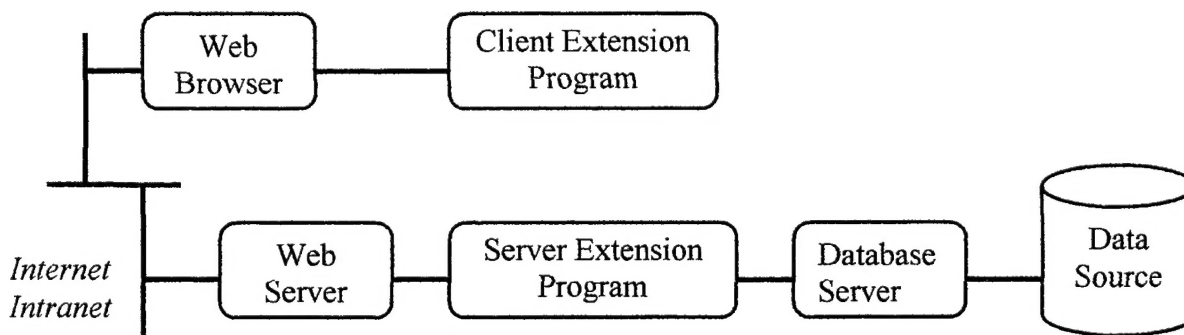


Fig.2.2 Database in a web-based system

2.1 References for Conventional database

1. Allen G. Taylor, "SQL for dummies," 3rd edition, IDG Books WorldWide, Inc., 1998, Foster City, CA.
2. Kevin Loney, Marlene Theriault, "Oracle 8i DBA handbook," Osborne/McGraw-Hill, 2000, San Francisco.
3. George Koch, Kevin Loney, "Oracle 8, the complete reference," Osborne, McGraw-Hill, 1997, San Francisco.

3. Neural networks for nonlinear pattern recognition

3.1 Nonlinear pattern recognition

Fig.3.1 illustrates the process of pattern recognition in a two-feature space; the feature distribution of three objects, A, B and C are shown. In Fig.3.1(a) classes A, B and C are linearly separable. Two linear decision hyperplanes, D1 and D2, can classify A, B and C very well with 100% accuracy. However, in Fig.3.1(b), the feature distribution of A, B and C are meshed, posing a problem for linear decision hyperplanes. The shaded areas show the decision error resulting from the nonlinear boundary condition of the feature space.

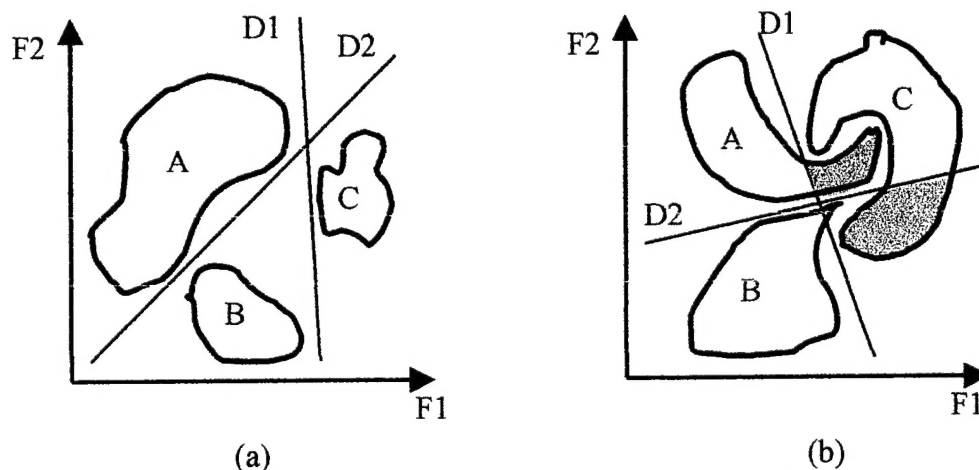


Fig.3.1 Object A, B and C in a two-feature space. (a) A, B and C are linearly separable by two hyperplanes, D1 and D2. (b) A, B and C are not linearly separable.

In the past, much research and development effort has attempted to make a computer perform pattern recognition^[1]. Many successful examples have been shown in the field of machine vision for production inspection and automation in a controlled environment. Well-defined linear algorithms have been developed to extract features from images and classify objects in real time with a computer. Specially designed electronic processors have been developed to speed up the recognition process. However, many practical problems, such as the processing of the 3D vision of an arbitrarily oriented object in a noisy, uncontrolled environment, are still very difficult for a computer to perform. These problems require nonlinear feature extraction and classification in a complex nonlinear feature space, and also demand the ability to adapt to a changing environment.

3.2 Artificial neural network, learning and parallel processing

Artificial neural network (NN) technology has been developed in the last two decades to attack these nonlinear pattern recognition problems. Artificial NN's are modeled after biological NN's, mimicking some of the recognition and deductive functions of the human brain^[2,3]. One of the most intriguing characteristics of an artificial

NN is its ability to modify dynamically its interconnection weights so that a predefined system behavior could be achieved^[4-9]. This interconnection weight modification process is termed as the learning of a neural network. Another very important property of an artificial NN is its parallel processing capability. After learning, a neural network can compare the input data against all stored patterns simultaneously, thus having very high processing speed. This is totally different from the serial read-and-compare processing of a computer-based pattern recognition.

Recently, specialized electronic NN processors and VLSI chips have been introduced into the commercial market^[10]. The number of parallel processing channels they can handle is limited because of the limited interconnection capability of electronic wires within an IC chip. For high-resolution object recognition, a large number of neurons will be required in order to achieve high parallelism processing.

Because of its inherent parallel processing ability, optical processors are particularly useful in feature extraction and information reduction at the preprocessing layer of a neural network where massive interconnection is usually required.

3.3 Neural networks for nonlinear processing

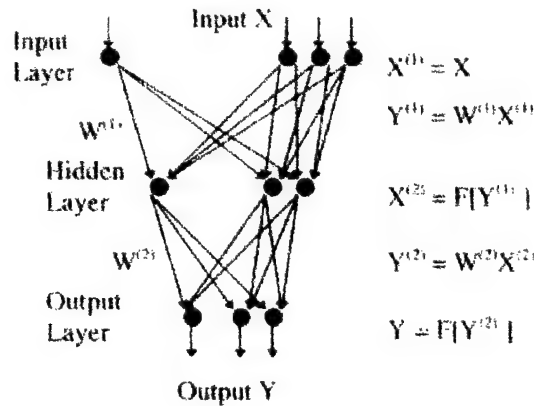


Fig.3.2 Operational flow of a multi-layer neural network. It consists of a series of matrix-vector multiplication and nonlinear thresholding.

As shown in Fig.3.2, a neural network consists of multiple layers of processors (neurons) that are interconnected. Input data enter the system at the input neuron layer. Each neuron performs weighted summation operation on all its inputs, compare the result with a predefined threshold, and then generates an output by using a nonlinear function, as shown below

$$y_i = f\left(\sum_j w_{ij}x_j - \theta_i\right) \quad (3.1)$$

where x_j ($j=1,2,\dots,N$) is the signal from the j^{th} neuron, w_{ij} is the weight associated with neuron i and j , θ_i is a bias value, and f is a nonlinear function, which can be a hard limit, a threshold logic, or a sigmoid function.

A multi-layer feed-forward NN can be represented as follows:

$$y_i^{(k)} = f \left[\sum_j w_{ij} y_j^{(k-1)} - \theta_i \right] \quad (3.2)$$

where $y_i^{(k)}$ is the i^{th} neuron output in the k^{th} layer, $f[]$ is the nonlinear transfer function of each neuron, w_{ij} is the interconnection weight between $y_i^{(k)}$ and $y_j^{(k-1)}$, and θ_i is the bias constant in each neuron. In the vector form, Eq.(3.2) can be expressed as

$$\mathbf{Y}^{(k)} = \mathbf{W}\mathbf{X}^{(k-1)} - \Theta \quad (3.3a)$$

$$\mathbf{X}^{(k)} = f[\mathbf{Y}^{(k)}] \quad (3.3b)$$

The learning process of a NN refers to the modification of its interconnection weight matrix \mathbf{W} based on the optimization of a predefined object function. The object function, for example, could be the difference between the practical output (which can be calculated from Eq.(3.3)) and the desired value. The interconnection matrix \mathbf{W} is optimized in such a way that makes the object function (the difference) reach its minimum. Many different learning algorithms have been proposed and tested^[11]. One of the most famous and very efficient algorithms is the BP (backward propagation) algorithm.

From Eq.(3.3) we can see that a single layer NN is a general matrix-vector product that can be trained to represent most linear transformations such as the Fourier transform, correlation, linear filtering, principal component analysis, and partial least-square analysis^[12]. However, a multi-layer neural network performs linear transformations within a single layer and nonlinear transfer functions between layers, as shown in Fig.3.2. Thus a multi-layer NN is a nonlinear processor that can be used to approximate any continuous nonlinear function with arbitrary desired accuracy. Hornik et al proved that a network with only one hidden layer of sigmoid neurons is enough to have universal approximation properties^[13]. One principal advantage of a multi-layer NN stems from the universal NN architecture that enables the network system to adapt to different environments through training by examples.

Fig.3.3 compares the classification capability of a linear classifier, a two-layer NN, and a three-layer nonlinear NN^[3]. Neither a single layer NN nor a linear classifier can separate two classes that have mashed features. In Fig.3.3 regions A and B represent the feature distribution of class A and B, respectively. The thin lines are the boundaries of the decision plane. As we can see, by proper training, the two- and three-layer NN's can cut through the mashed areas of two classes and form a nonlinear boundary to separate the two classes. The misclassification error rate can be reduced further by multi-layer NN, since the decision plane would be more adaptive.

A neural network can learn from the training examples to associate features of input patterns with an optimal (desired) output result. The adaptive learning, massive interconnection, and nonlinear classification capabilities of NN's make them generally more robust to noise and distortion.


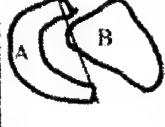
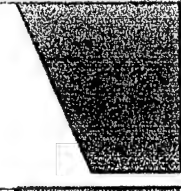






Structure	Type of Decision Regions	Class with Mashed Regions	Most General Region Shapes
 Linear, Single Layer	Half-Plane Bounded by Hyperplane		
 Two Layer	Convex Open or Closed Regions		
 Three Layer	Arbitrary Limited by Number of Nodes		

Fig.3.3 Comparison of a linear classifier, a two-layer NN and a three-layer NN. The thin lines represent the decision plane boundary.

3.4 Holographic optical neural network system

3.4.1 Massive interconnection requirement

The operation of a one-layer 2 dimensional NN can be expressed in the following matrix-vector product form:

$$y_{ij} = f \left[\sum_{l=1}^N \sum_{k=1}^N (T_{ij})_{lk} x_{lk} \right] \quad i, j=1, 2, \dots, M \quad (3.4)$$

where y_{ij} and x_{lk} are the states of the ij^{th} output neuron and the lk^{th} input neuron, respectively, $(T_{ij})_{lk}$ is the set of interconnection weights between these pairs of neurons, f is the nonlinear threshold function, and $N \times N$ and $M \times M$ are the number of input and output neurons in the network, respectively.

Suppose $M=N$. Then from Eq.(3.4) we see that, an $N \times N$ neural network has N^4 interconnections. For a 512×512 NN, $(512)^4$ or 7×10^{10} interconnections must be implemented. It is far beyond state-of-the-art VLSI technology to fabricate a chip or a board of so many interconnections to perform parallel pattern recognition.

Optical technologies, by virtual of their inherent 3D global interconnection capability, are good candidates for implementing this massive interconnection and parallel processing in the first layer of a multi-layer NN. Photorefractive crystals have the potential of dynamic learning^[14]. Passive holographic materials, such as dichromated gelatin, silver halide, Dupont and Polaroid photopolymers, are well developed. They possess the properties of high resolution (>5000 line pair/mm), high refractive index

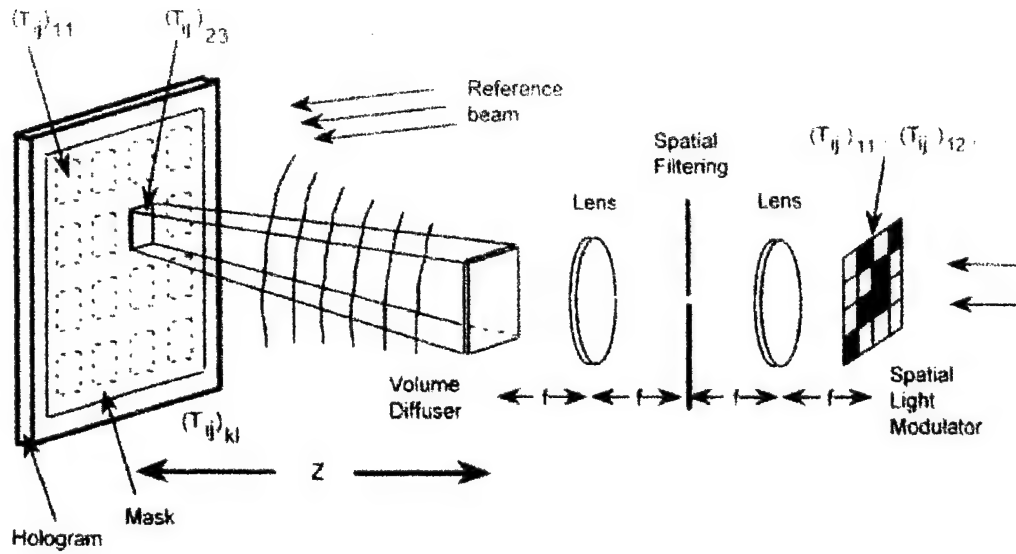
change, large recording area and low cost. These holographic materials offer an ideal means of a massively parallel 3D interconnection for a large scale NN implementations [15,16]. The interconnections in the first layer may be fixed as a static associative memory by use of holographic implementations. The subsequent layers are much smaller than the first layer, and then can be implemented by an electronic NN for adaptive training and nonlinear recognition.

3.4.2 Implementation of a holographic optical neural network

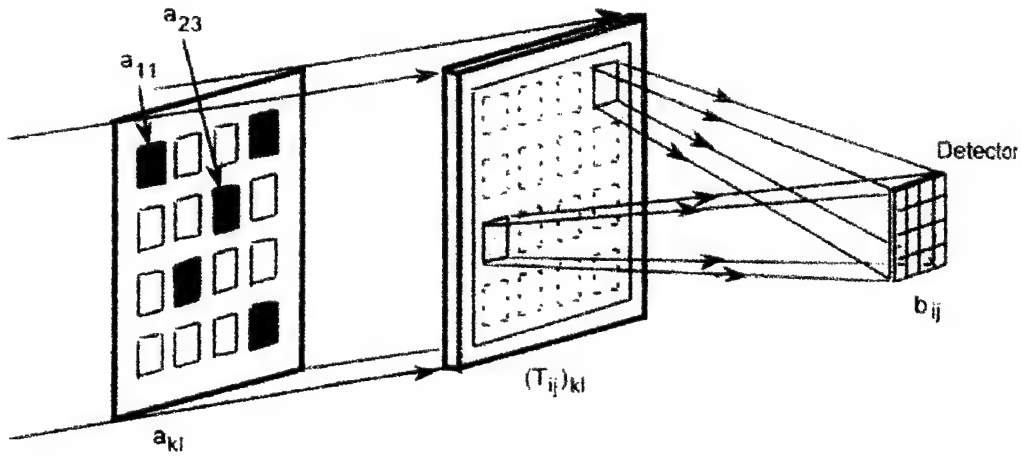
In this section we show an example implementation [16-20] of a holographic optical neural network with high-density interconnection capacity, developed by Physical Optics Corp., Torrance, CA. Fig.3.4 shows the schematic diagram of the setup for recording an N^4 interconnection weight matrix.

In the setup shown in Fig.3.4, a collimated beam illuminates an SLM at one particular incidence angle (near normal, for example). A diffuser, placed in the image plane of a $4f$ imaging optical system, spreads the incident pattern (modulated by the SLM) over a wide angular range. The $4f$ imaging system faithfully images the SLM pixel array onto the diffuser. A hologram plate with a mask is placed at a proper distance Z (which depends on the desired recording size and the spread angle of the diffuser) behind the diffuser. With an additional reference beam, the pattern from the SLM (T_{ij}) will be holographically recorded in an element of the array. By changing the SLM pattern and moving the mask along both horizontal and vertical directions, one can fabricate an N^4 interconnection weight matrix [ie $(T_{ij})_{jk}$]. The weight information can be coded by varying the ON time of each SLM pixels. Thus, the longer the two-beam exposure, the stronger the grating strength and the higher the diffraction efficiency. The key element in the recording process is the design of the diffuser, because its characteristics, such as speckle size, directionality, and uniformity are crucial to the performance of the N^4 holographic element array.

In the reconstruction process (see the bottom half of Fig.3.4), an encoded reference beam (with pixel a_{kl}), which is conjugated to the reference beam in the recording process and represents the input information, illuminates the holographic array. This second reference beam can be realized by an SLM or can originate from an array of laser diodes with collimating lenses. A photo-detector array that has the same packing density and pixel size as the SLM used in the recording process is placed at the diffuser position. The beams diffracted from the holographic matrix elements are directed to the photo-detector array exactly.



Recording setup for an N^4 holographic IWM with a diffuser.



Reconstruction of an N^4 holographic IWM.

Fig. 3.4 Recording and reconstruction of a large scale holographic optical neural network.

Because the holographic array consists of many matrix elements, the outputs of these holographic elements add up pixel by pixel in the photo-detector array. Thus the resulting signals detected by one of the detectors can be described as

$$b_{ij} = \sum_{k,l} (T_{ij})_{kl} a_{kl} \quad (3.5)$$

This equation describes the parallel interconnections of an $N \times N$ input array to an $N \times N$ output array through an N^4 holographic interconnection weight matrix T .

3.4.3 Pattern recognition

Many successful pattern recognition operations based on the holographic optical neural network described above have been demonstrated^[16-20]. The complete system is of course an optical-electronic hybrid implementation. Input images are first electronically preprocessed, including smoothing, normalization and edge enhancement. The preprocessed signals are then sent to the optical NN for training and object identification. The input neurons are grouped to represent the feature windows from the preprocessing stage. A set of examples is used to train the NN to produce the correct responses from the output neurons such that only one output neuron responds high and the rest low for certain images. The computation time for training the system is less than 15 seconds. The examined patterns include pictures of military vehicles like planes, tanks and helicopters. Based on real-time pattern recognition, an automatic tracking system has also been developed^[18].

3.5 Comments

Due to its inherent 3D global interconnection and parallel processing capabilities, optical technology has great potential for the implementation of large-scale neural networks for nonlinear pattern recognition. So far the main difficulty in building an optical information processing system is the lack of proper optical-electronic (OE) and electronic-optical (EO) devices, and optical recording materials. The desired OE and EO (such as CCD and SLM) must have high speed, high sensitivity, and high resolution. The desired optical materials should be of large dynamic range, high resolution, high response speed and/or long lifetime. Also, much attention should be paid to the design of the interface between an optical information processing system and its associated electronic unit. The interface must be of high enough speed so that the high-speed advantage of the optical implementation can be fully leveraged.

3.6 References for neural networks for nonlinear pattern recognition

1. D. Marr, "Vision", Freeman, San Francisco, 1982.
2. D. E. Rumelhart and J. L. McClelland, eds, "Parallel distributed processing," MIT, Cambridge, MA, Vol.1, 1986.
3. R. P. Lippmann, "An introduction to computing with neural nets," IEEE Trans. Acoust. Speech Signal Process., Vol.4, pp.4-22, 1987.
4. H. A. Castro, S. M. Tam, and M. Holler, "Implementation and performance of an analog nonvolatile neural network," Analog Integrated Circuits and Signal Processing, vol.4, no.2, pp.97-113, 1993.
5. V. Forman et al., "Multisensor target recognition system (MUSTRS)," in Proceedings of the Third Automatic Target Recognition System and Technology Conference, GACIAC PR 93-01, 1993.
6. Bowman, "Artificial neural network approaches to target recognition," in Proceedings of IEEE Digital Avionic System Conference (IEEE, New York), pp.847-857, 1988.

7. N. Farhat, "Smart sensing and recognition based on models of neural networks," Final report to U.S. Army Research Office (U.S. Army Research Office), 1990.
8. J. Llinas, and E. Waltz, "Multisensor data fusion," Artech House, Norwood, MA, 1990.
9. Marchette and C. Priebe, "An application of neural networks to a data fusion problem," in Proceedings of the Data Fusion Symposium, Vol.1, pp.230-235, 1987.
10. Gert Cauwenberghs, Magdy A. Bayoumi, "Learning on silicon : adaptive VLSI neural systems," Boston, Kluwer Academic, 1999.
11. Martin Anthony and Peter L. Bartlett, "Neural network learning: theoretical foundations," Cambridge, U.K. ; New York, NY : Cambridge University Press, 1999.
12. T. Lu, J. M. Lerner, "Spectroscopy and hybrid neural network analysis," Proc. IEEE, Vol.84, pp.895-905, 1996.
13. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feed-forward neural networks are universal approximators," Neural Networks, Vol.2, pp.359-366, 1989.
14. Psaltis, D. Brady, and K. Wagner, "Adaptive optical neural networks using photorefractive crystals," Appl. Opt., Vol. 27, p.1752, 1988.
15. H. J. Caulfield, "Parallel N^4 weighted optical interconnects," Appl. Opt., Vol.26, p.4039, 1987.
16. Lin, "Practical realization of N^4 optical interconnections," Appl. Opt., Vol.29, pp.5226-5227, 1990.
17. T. Lu, F. Lin, "Experimental demonstration of large-scale holographic optical neural network," in Proceedings of the International Joint Conference on Neural Networks, Vol.1, pp.535-540, 1991.
18. A. Kostrzewski, T. Lu, H. Chou, F. Lin, "Portable holographic neural system for distortion-invariant pattern recognition and tracking in controlled environment," in Optical Pattern Recognition □, D. Casasent and T. H. Chao, eds, Proc. SPIE, Vol.1701, pp.298-307, 1992.
19. T. Lu, D. Mintzer, A. Kostrzewski, F. Lin, "Compact holographic optical neural network system for real-time pattern recognition," Opt. Eng, Vol.35, pp.2122-2131, 1996.
20. T. Lu, A. Kostrzewski, H. Chou, S. Wu, and F. Lin, "Performance evaluation of a holographic optical neural network (HONN) system," in Photonics for Computers, Neural Networks, and Memories, S. T. Kowel, W. J. Miceli, and J. A. Neff, eds, Proc. SPIE, Vol. 1773, pp.46-52, 1992.

4. Digital image associative search

4.1 Retrieval of image information

Rapid advancement in digital storage and processing technology has created a significant impact on the way information is stored and utilized. Early computers were only capable of storing and processing text based information, but the rapid development of digital technology and high-speed networks, and the proliferation of the World Wide Web has significantly changed this situation. Now, a large portion of the information used daily by individuals and organizations consists of images. Presently images are being generated at an ever-increasing rate by sources such as defense and civilian satellites, scientific experiments, space telescopes, medical imaging and home entertainment systems. These sources are so diverse and large in number that it is difficult to enumerate all of them. Context-based image retrieval provides a tool for efficient retrieval of images from large image repositories.

4.2 Content based image retrieval

Content Based Image Retrieval or CBIR is a broad term, which can be defined as *“using a complete or a partial image to index through a database of images to compare certain features and retrieve one or more images, which have minimum distance from the query image”*.

The importance of CBIR can be realized from the fact that images are being generated at an ever-increasing rate by sources such as defense and civilian satellites, military reconnaissance and surveillance flights, fingerprinting, scientific experiments, biomedical imaging, and home entertainment systems. For example, the Hubble space telescope daily generates terabytes of image data. A CBIR system is required to use information from these image repositories effectively and efficiently, allowing users to retrieve relevant images based on image contents. Application areas in which CBIR is a principal activity are numerous and diverse [3]:

- Art galleries and museum management,
- Architectural and engineering design,
- Interior design,
- Remote sensing and management of earth resources,
- Geographic information systems,
- Scientific database management,
- Weather forecasting,
- Retailing,
- Fabric and fashion design,
- Trademark and copyright database management,
- Law enforcement and criminal investigation,
- Picture archiving and communication systems, and
- Document storage and processing

4.3 Approaches

With the advancement in the digital data processing technology, CBIR has taken a different direction lately. Previously, image contents were modeled as a set of attributes extracted manually and managed within the framework of conventional database-management systems. Queries were specified using these attributes. Attribute-based representation of images requires a high level of image abstraction. Generally, the higher the level of abstraction, the lesser is the scope for posing ad hoc queries to the image database. Attribute-based retrieval has been advocated and advanced primarily by database researchers.

Lately the new direction by CBIR depends on an integrated feature-extraction/object recognition subsystem to overcome the limitation of attribute based retrieval. This subsystem automates the feature-extraction and object-recognition task that occurs when the image is inserted into the database. Though these approaches are computationally expensive as compared to attribute-based approaches, fast digital processing technology has made these approaches possible.

CBIR research draws upon ideas from areas such as knowledge-based systems, cognitive science, user modeling, computer graphics, image processing, pattern recognition, database management systems and information retrieval. This confluence of ideas has culminated in the introduction of novel image representations and data models, efficient and robust query-processing algorithms, intelligent query interfaces, and domain-independent system architectures.

4.4 Query classes

Generic query classes facilitate CBIR through retrieving by:

- Color,
- Texture,
- Sketch,
- Shape,
- Volume,
- Spatial constraints,
- Browsing,
- Objective attributes,
- Subjective attributes,
- Motion,
- Text, and
- Domain concepts

A CBIR system featuring all these query classes would have reasonable generality for dealing with diverse applications.

Color and *texture* queries let users select images containing objects specified accordingly. Retrieval by sketch lets users outline an image and then retrieves a like image from the database. This class can be thought of as retrieving images by matching the dominant edges. The *shapes* class of queries has a counterpart in 3D images referred to as *Retrieval by volume*. The spatial constraints category deals with a class of queries

based on spatial and topological relationships among the objects in an image. These relationships may span a broad spectrum ranging from directional relationships to adjacency, overlap, and containment involving pair of objects or multiple objects.

Retrieval by browsing is performed when users are vague about their retrieval needs or are unfamiliar with the structure and types of information available in the image database. The *objective attributes* query uses attributes like the date of image acquisition or the number of bedrooms in a residential floor-plan image, and is similar to Structured Query Language retrieval in conventional databases. Retrieval is based on exact match of attribute values. In contrast, a *subjective attributes* query is characterized by the presence of attributes that may be interpreted differently by each user. For example, in mug-shot database, the attribute *eyebrow shape* assumes one of three values: arched, normal or straight. One user may assign the normal value for the eyebrow shape, while another may interpret the value as arched.

Retrieval by motion facilitates retrieving spatiotemporal image sequences depicting a domain phenomenon that varies in time or geographic space. Some applications require retrieving images based on associated text. Such a need is modeled by retrieving by text. Note that processing this query involves natural language processing and information retrieval techniques.

The above mentioned query classes can be used as fundamental operators in formulating a class of complex queries referred to as *Retrieval by Domain Concepts*. An example of this is "Retrieve images of snow-covered mountains".

Not all the above generic query classes are necessary, however, for a given image retrieval application. For example, a real estate marketing application may require only retrieval by browsing, objective attributes, shape, and spatial constraints. In such a case, the a priori feature-extraction approach helps generate an application-specific system from a generic one by retaining only the necessary classes.

4.5 Using color to retrieve image

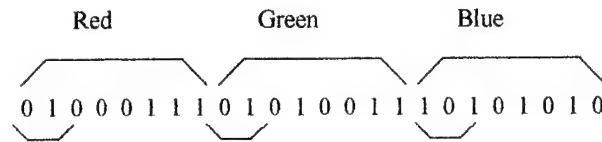
Taking color as an instance of the query classes, we will in this section explain how an image is retrieved from an image database, using a technique based on color [2]. Though there are a number of techniques utilized in this context, one important approach is the use of *Color histograms*. Color histograms themselves have certain limitations and so they are not usually utilized as stand-alone image retrieval techniques; rather they are typically utilized as an integral part of a more sophisticated technique. A number of major CBIR systems, which are available online over the World Wide Web, utilize this technique (See Section 4.14).

4.6 Color histogram

Color histograms provide us a very general picture of the colors in an image. They provide us with a vector showing the color content of the image at a coarse level.

Consider an image having certain number of pixels say M pixels. For simplicity it is assumed that all the images both in the database and the query image have the same number of pixels. If we consider 24-bit image then the image possibly contains 16 million colors. This makes the computation fairly complex. Instead, the color space is

discretized into a small number n colors so that the vector is manageable. This is done by taking few most significant bits of each primary color of each pixel. For example, consider a pixel representation:



Taking only the first 2 bits of each primary color gives us a total of 6 bits. Using these 6 bits a total of $2^6 = 64$ colors can be represented. These colors are called buckets. For each pixel, the bucket is computed into which the pixel falls. In this way the total number of pixels falling into each bucket are counted. Accumulating these counts into a vector give us the color histogram. It can be represented as:

$$H_I = \langle h_1, h_2, h_3, \dots, h_n \rangle$$

Where H_I is the color histogram thus obtained, and $h_1, h_2, h_3, \dots, h_n$ are the number of pixels falling into each bucket.

Usually the color histogram of each image is calculated when the image is being inserted into the database. The histogram is stored along with the image in the database so that it can be retrieved and processed easily whenever required.

Any image query is processed by computing the distances between the query image and different images in the database, and then picking up one or a few images having the minimum distance. Typically color histograms are compared using the sum of squared differences (L_2 -distance) or the sum of absolute value of differences (L_1 -distance). So, the most similar image I would be the image I' minimizing

$$\| H_I - H_{I'} \| = \sum (H_I[j] - H_{I'}[j])^2,$$

for the L_2 -distance, or

$$\| H_I - H_{I'} \| = \sum | H_I[j] - H_{I'}[j] |,$$

for the L_1 -distance. Note that it is assumed that differences are weighted evenly across different buckets for simplicity.

4.7 Advantages of color histograms

Color histograms bear certain advantages over conventional techniques:

- They are computationally trivial.
- Small change in the camera viewpoint tend not to effect color histograms, and
- Different objects often have distinctive color histograms.

4.8 Limitations of color histograms

In addition to the above mentioned advantages, the color histograms have certain limitations:

- They do not take into account any spatial information present in the image and they only quantify the colors present. If two different images have similar color content, they will be regarded as similar if this technique is used.
- Color histograms are sensitive to compression artifacts and changes to overall image brightness.

4.9 Color coherence vectors (CCV)

The above-mentioned limitations of color histograms have been addressed by another technique called *Color Coherence Vectors*. This technique is an enhancement of the color histograms as explained below.

Color coherence is defined as the degree to which pixels of a given color are members of large similarly colored regions. These significant regions can be regarded as *coherent regions*, and are of significant importance in characterizing images.

The coherence measure classifies the pixels as either coherent or incoherent. Coherent pixels are a part of some sizable contiguous region, while incoherent pixels are not. A *color coherence vector (CCV)* represents this classification for each color in the image. CCV's prevent coherent pixels in one image from matching incoherent pixels in another. This allows fine distinction that cannot be made with color histograms.

4.10 CCV computation

The initial stage of CCV computation is similar to the computation of a color histogram. In the first step, the image is slightly blurred by replacing pixel values with the average value in a small local neighborhood (including the 8 adjacent pixels). This eliminates small variations between neighboring pixels. The color space is then, discretized such that there are only n distinct colors in the image.

The next step is to classify the pixels within a given color bucket as either coherent or incoherent. A coherent pixel is part of a large group of pixels of the same color, while incoherent pixel is not. Pixel groups are determined by connected components. A connected component C is a maximal set of pixels such that for any two pixels $p, p' \in C$, there is a path in C between p and p' . (Formally, a path in C is a sequence of pixels $p = p_1, p_2, \dots, p_n = p'$ such that each pixel p_i is in C and any two sequential pixels p_i, p_{i+1} are adjacent to each other. Any two pixels are adjacent if one pixel is among the eight closest neighbors of the other; in other words, diagonal neighbors are included.) Note that the connected components are computed within a given discretized color bucket. This effectively segments the image based on the discretized colorspace.

Connected components can be computed in linear time. When this is done, each pixel belongs to exactly one connected component. The pixels are, then, classified as either coherent or incoherent depending on the size in pixels of its connected components. A pixel is coherent if the size of its connected components exceeds a fixed value τ ; otherwise, the pixel is incoherent.

For a given discretized color, some of the pixels with that color will be coherent and some will be incoherent. Let us call the number of coherent pixels of the j 'th

discretized color α_j and the number of incoherent pixels β_j . Clearly the total number of pixels with that color is $\alpha_j + \beta_j$, and so a color histogram would summarize the image as

$$\langle \alpha_1 + \beta_1, \dots, \alpha_n + \beta_n \rangle.$$

Instead, for each color the pair is computed

$$(\alpha_j, \beta_j)$$

which is called the coherence pair for the j 'th color. The *color coherence vector* for the image consists of

$$\langle (\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n) \rangle.$$

This forms a vector of coherence pair one for each discretized color.

Usually τ is determined to be a certain percentage of the image. For example, in an image containing 38,976 pixels, a τ of 300 pixels would be less than 1% of the total number of pixels in the image.

4.11 Comparing CCV's

Consider to images I and I' , together with their CCV's G_I and $G_{I'}$, and let the number of coherent pixels in color bucket j be α_j (for I) and α'_j (for I'). Similarly let the number incoherent pixels be β_j and β'_j . So

$$G_I = \langle (\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n) \rangle$$

$$G_{I'} = \langle (\alpha'_1, \beta'_1), \dots, (\alpha'_n, \beta'_n) \rangle$$

Color histograms will compute the difference between I and I' as

$$\Delta_H = \sum |(\alpha_j + \beta_j) - (\alpha'_j + \beta'_j)|$$

This equation can be rearranged to get the following:

$$\Delta_G = \sum |(\alpha_j - \alpha'_j)| + |(\beta_j - \beta'_j)|$$

These two equations show that CCV's create a finer distinction than color histograms. A given color bucket j can contain the same number of pixels in I as in I' , i.e.

$$\alpha_j + \beta_j = \alpha'_j + \beta'_j$$

but these pixels may be entirely coherent in I and entirely incoherent in I' . In this case, $\beta_j = \alpha'_j = 0$, and while $\Delta_H = 0$, Δ_G will be large.

In general $\Delta_H \leq \Delta_G$. This is true even if we used squared differences instead of absolute differences in the definitions of Δ_H and Δ_G .

4.12 Computational efficiency

There are two phases to the computation involved in querying an image database. First, when an image is inserted into the database, a CCV must be computed. Second, when the database is queried, some number of the most similar images must be retrieved. Most methods for content-based indexing include these distinct phases. For both color histograms and CCV's, these phases are implemented in linear time.

According to the experimental results given in [2], on a 50MHz SPARCstation 20, color histograms can be computed at 67 images per second, while CCV's can be computed at 5 images per second. Using color histograms, 21,940 comparisons can be performed per second, while with CCV's 7,746 can be performed per second. The images used for the benchmarking were 232 X 168.

4.13 References for digital image associative search

1. U. Gargi and R. Kasturi, "Image Database Querying Using a Multi-Scale Localized Color Representation", in IEEE Workshop on Content-Based Access of Image and Video Libraries, June 22, 1999.
2. G. Pass, R. Zabih and J. Miller, "Comparing Images Using Color Coherence Vectors", in ACM International Conference on Multimedia, November 1996.
3. V. Gudivada and V. Raghavan, "Content-Based Image Retrieval Systems", in IEEE Computer, 28(9):18-22, September 1995.
4. M. Flickner *et al*, "Query by Image and Video Content: The QBIC System", in IEEE Computer, 28(9):23-32, September 1995.
5. R. Srihari, "Automatic Indexing and Content-Based Retrieval of Captioned Images", in IEEE Computer, 28(9):49-56, September 1995.
6. W. Hsu, Chau T. S. and Pung H. K., "An Integrated Color Spatial Approach to Content-Based Image Retrieval", in ACM Multimedia Conference, November 5, 1995.
7. J. Huang, S. Kumar, M. Mitra, W. Zhu and R. Zabih, "Image Indexing Using Color Correlograms", in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1997.
8. Berman and L. Shapiro, "Efficient Content-Based Retrieval: Experimental Results", in IEEE Workshop on Content-Based Access of Image and Video Libraries, June 22, 1999.
9. Berman and L. Shapiro, "Efficient image retrieval with multiple distance measures", in Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases, February, 1997.
10. Berman and L. Shapiro, "Selecting good keys for triangle-inequality-based pruning algorithms", in IEEE International Workshop on Content-Based Access of Image and Video Databases, January 1998.
11. Berman and L. Shapiro, "A Flexible Image Database System for Content-Based Retrieval", in 17th International Conference on Pattern Recognition, August 1998.
12. Berman and L. Shapiro, "Triangle-Inequality-Based Pruning Algorithms and with Triangle Tries", in Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases, January 1999.
13. J. Barros, J. French, W. Martin, P. Kelley, and M. Cannon, "Using the triangle inequality to reduce the number of comparisons required for similarity-based retrieval", in IS&T/SPIE-Storage and Retrieval for Still Image and Video Databases, Volume IV, 1996.
14. D. Forsyth, J. Malik, M. Fleck, H. Greenspan, T. Leung, S. Belongie, C. Carson and C. Bergler, "Finding pictures of objects in large collections of images", in Proceedings of the 2nd International Workshop on Object Representation in Computer Vision, 1996.
15. T. Kato, T. Kurita, N. Otsu and K. Hirata, "A sketch retrieval method for a full color image databases", in 11th International Conference on Pattern Recognition, 1992.
16. Bimbo, M. Campanai and P. Nesi, "3D visual query language for image databases", in Journal of Visual Language and Computing, Vol 3, 1992.

18. Pentland, R. Picard and S. Sclaroff, "Photobook: tools for content-based manipulation of image databases", Technical Report, Volume 255, MIT, Media Lab., 1993.
- V. Ogle and M. Stonebraker, "Chabot: Retrieval from a relational database of images", in IEEE Computer, 28(9):40-48, September 1995.

4.14 Web Sites for Content-Based Image Retrieval

- [1] <http://rvl2.ecn.purdue.edu/~cbirdev/WWW/CBIRmain.html>
Content-Based Image Retrieval for Medical Image Databases
- [2] http://www.research.ibm.com/networked_data_systems/9auc/whitepaper/paper.html
Scalable Content-Based Retrieval from Distributed Image/Video Databases
- [3] <http://maya.ece.ucsb.edu/Netra/>
NETRA: A Content-Based Image Retrieval System
- [4] <http://www.cs.umd.edu/~fatma/abstract.html>
Comparison of Content Based Image Retrieval Systems
- [5] <http://www-db.stanford.edu/~wangz/project/imsearch/>
Content-based Image Retrieval Project
- [6] <http://www.jtap.ac.uk/projects/jtap-564.html>
Report on Content-based Image Retrieval
Dr John Eakins, Institute for Image Data Research
University of Northumbria at Newcastle, Ellison Place, Newcastle upon Tyne NE1 8ST
- [7] http://www.research.microsoft.com/mlp/Pages/proj_retrieve.htm
Content-based Image Retrieval—research at Microsoft
- [8] <http://elib.cs.berkeley.edu/seminar/1999/19991029.html>
SPIRE: A Progressive Content-Based SPatial Image Retrieval Engine
- [9] <http://www.unn.ac.uk/iidr/research/cbir/report.html>
Content-based Image Retrieval. A report to the JISC Technology Applications Programme
John P Eakins and Margaret E Graham
Institute for Image Data Research, University of Northumbria at Newcastle, January 1999
- [10] http://www-rocq.inria.fr/imedia/index_UK.html
The IMEDIA Project. Image and multimedia indexing, browsing and retrieval

- [11] <http://www.cs.washington.edu/research/imagedatabase/>
University of Washington. Efficient Content-Based Image Retrieval
- [12] <http://elib.cs.berkeley.edu/vision.html>
Computer Vision Research. UC Berkeley Computer Vision Group.
- [13] <http://www.cis.hut.fi/picsom/publications.html>
Demonstration of PicSOM system for content-based image retrieval
- [14] <http://maya.ctr.columbia.edu:8080/>
IBM-NASA Satellite Image Explorer
- [15] <http://cui.unige.ch/~muellerh/contentBasedImageRetrieval.html>
Content based Image Retrieval
- [16] <http://www.dlib.org/dlib/february97/columbia/02chang.html>
Finding Images/Video in Large Archives. Columbia's Content-Based Visual Query Project
- [17] <http://www.cs.bris.ac.uk/Tools/Reports/Abstracts/1998-wood.html>
Iterative Refinement by Relevance Feedback in Content-Based Digital Image Retrieval
- [18] http://meru.cecs.missouri.edu/mm_seminar/cont_ret.html
Content-Based Image Retrieval by Using Color and Texture Information
- [19] http://www.ctr.columbia.edu/~jrsmith/html/pubs/TNT/color_tech_1.html
Tools and Techniques for Color Image Retrieval
- [20] <http://cui.unige.ch/~viper/>
Computer Vision Group
Department of Computer Science
University of Geneva, Switzerland
- [21] <http://www.ctr.columbia.edu/VisualSEEK/>
VisualSEEK is a Web tool for searching for images and videos. VisualSEEK allows the user to make queries using visual features. The demo system currently supplies 3,200 (and now also 12,000) color images and videos. Queries may be conducted by sketching the layout of color regions, by providing the URL of a seed image or by using instances of prior matches. Image and ATV Lab of Columbia University.
- [22] <http://www.attrasoft.com/imagefinder42/imageretrieval.html>
Image retrieval, classification, software, etc. at Attrasoft.

- [23] http://www.brunel.ac.uk/depts/ee/Research_Programme/NN/res-colcbir.html
Colour Content-Based Image Retrieval
- [24] <http://www.ee.umd.edu/medlab/papers/draft1c/draft1c.html>
Heuristic Similarity Measure Characterization for Content-Based Image Retrieval
Wilbur S. Peng - Nicholas DeClaris
Medical Informatics & Computational Intelligence Laboratory
University of Maryland at College Park
- [25] <http://vivaldi.ece.ucsb.edu/projects/ipinadl/ipadl.html>
Efficient Management of Image and Video Data. Alexandria Digital Library Project
- [26] <http://public.lanl.gov/kelly/CANDID/index.shtml>
The CANDID Project--the Comparison Algorithm for Navigating Digital Image Databases. Los Alamos National Laboratory, Operated by the University of California for the US Department of Energy.
- [27] <http://http.cs.berkeley.edu/~ginger/chabot.html>
Chabot: Retrieval from a Relational Database of Images, UC Berkeley
- [28] <http://www.tzi.de/bv/ImageMinerhtml/>
ImageMinerTM combines methods and techniques of computer vision and knowledge representation in a novel way in order to automatically generate textual content descriptions of images.
- [29] <http://www-white.media.mit.edu/~roz/res-diglib.html>
Research in Video and Image Libraries: Browsing, Retrieval, Annotation.
- [30] <http://www.ifp.uiuc.edu/~yrui/html/research.html>
Content-based Texture Image Retrieval with Relevance Feedback. Personal website.
- [31] <http://unicorn.comp.nus.edu.sg/~mmir/>
Visual property-based search engine for image retrieval.
- [32] <http://www.brunel.ac.uk/~eepgmas/report/report.html>
On the use of colour in content based image retrieval
Seaborn, M., submitted for 1st year report of PhD
Department of Electronic and Electrical Engineering
Brunel University
- [33] http://eassun.eas.gu.edu.au/res_retri.htm
Content Based Image Retrieval
- [34] <http://www.gl.umbc.edu/~kwater2/paper/blobsketch.htm>
BlobSketch: A Web Interface for Content-based Image Retrieval
Kimberly G. Waters, University of Maryland, Baltimore County

- [35] <http://www.jtap.ac.uk/reports/htm/jtap-039.html>
Content-based Image Retrieval
John Eakins and Margaret Graham, University of Northumbria at Newcastle
- [36] <http://www.ctr.columbia.edu/MetaSEEk/MetaSEEk.html>
(Paper) A Content-Based Image Meta-Search Engine using Relevance Feedback
Ana B. Benitez, Mandis Beigi, and Shih-Fu Chang
Department of Electrical Engineering & New Media Technology Center, Columbia University
- [37] <http://www.psc.edu/research/abstracts/becich.html>
Content Based Image Retrieval and Pathology Image Classification Image Processing
- [38] <http://www.unn.ac.uk/iidr/CBIR/cbir.html>
The Institute for Image Data Research.
- [39] <http://icg.cityu.edu.hk/KENT/ImageDB.htm>
Content-based Image Database Retrieval By Generalized Complex Moments
- [40] <http://compass.itc.it/links.html>
Computer aided search system.
- [41] <http://www.lic.gov.uk/awards/ir-curpj.html>
This list includes details of current projects supported by the Library and Information Commission through its Information Retrieval research programme.

5. Volume holographic associative memory and other optical correlators

5.1 Volume holographic associative memory

This approach has been investigated intensively in the past few years because of renewed interest in the commercialization of holographic data storage. For detailed information on volume holographic associative memory, please refer to "Xiaochun Li, Fedor Dimov, William Phillips, "Low-cost optical search of digital holographic storage systems", Final report to Air Force Rome Laboratories. Contract No. F30602-97-C-0343. Nov, 2000."

5.2 References for volume holographic associative memory

The following papers have more detailed information on volume holographic associative memory:

1. Xiaochun Li, Fedor Dimov, William Phillips, "Low-cost optical search of digital holographic storage systems", Final report to Air Force Rome Laboratories. Contract No. F30602-97-C-0343. Oct., 2000.
2. Xiaochun Li, Fedor Dimov, William Phillips, Lambertus Hesselink, Robert McLeod, "Parallel associative search by use of a volume holographic memory," Applied Imagery Pattern Recognition (AIPR) 2000, Washington DC, Oct.16-18, 2000.
3. Geoffrey W. Burr, Sebastian Kobras, Holger Hanssen, Hans Coufal, "Content-addressable data storage by use of volume holograms," App. Opt., Vol.38, No.32, pp.6779-6784, 1999.
4. J. Ashley, M.-P. Bernal, G. W. Burr, H. Coufal, H. Guenther, J. A. Hoffnagle, C. M. Jefferson, B. Marcus, R. M. Macfarlane, R. M. Shelby, G. T. Sincerbox, "Holographic data storage," IBM Journal of Research & development, Vol.44, No.3, 1999. (Download from <http://www.research.ibm.com/journal/rd/443/ashley.html>).
5. Pericles A. Mitkas, George A. Betzos, Sakellaris Mailis, Nikos A. Vainos, "Characterization of associative recall in a volume holographic database system for multimedia applications," SPIE, Vol.3388, pp.198-208, April 1998.
6. Jian Fu, Marius P. Schamschula, H. John Caulfield, "Optical parallel database management system for page oriented holographic memories," Optics Express, Vol.5, No.12, pp.273-285, Dec. 1999.

5.3 Other optical correlators

Basically optical correlators can be divided into 2 classes: coherent and incoherent correlators. In most cases optical correlators calculate the correlation (inner product) between two images as their similarity measure.

The best-known optical correlator is the VanderLugt planar holographic correlator^[1-4]. The recorded hologram serves as the matched filter. The resulting signal-to-noise ratio (SNR) is usually high. The main problems of the VanderLugt correlator include the off-line filter recording and the need for a separate filter for each specific image pattern. Joint transform correlator (JTC)^[2-8] overcomes the problem of off-line

recording of a VanderLugt correlator by jointly transforming the input patterns and recording the resultant hologram. However, due to the involvement of nonlinear response, the SNR resulting from a JTC is usually smaller than that of a VanderLugt correlator^[9]. Angle-multiplexed volume holographic correlator^[10, 11] can search into multiple images simultaneously which are recorded in one common location, thus having very high processing speed. Volume holographic correlator is further characterized by its limited shift invariance when comparing with its planar holographic counterpart. This limitation is especially rigorous in the Bragg selective direction of volume holograms. This system lends itself best to search of relational databases, where information can have a fixed position on the stored record. The multistage holographic optical random access memory (HORAM), proposed by Liu et al^[12, 13] recently, makes use of binary and/or holographic optical devices to create multiple replicas of the input image spectrum patterns. These replicas readout multiple recorded holograms simultaneously, thus increasing the degree of parallelism. Now the main problem related to HORAM comes from the limited performance of the binary/holographic beam splitters. Diffraction efficiency and uniformity are the main concerns.

The basic incoherent optical correlator is the shadow-casting correlator^[14, 15]. It has very large 2D shift invariance. The system is compact and low cost. Diffraction arising from the high frequency component of input images will reduce the signal-to-noise ratio. To maintain reasonable SNR, the distance between the two input images should be limited, which in turn limits the achievable shift invariance amount. Using both positive and negative cycling-encoding^[14], the shadow-casting correlator can also calculate the absolute difference between two input images as their similarity measure. The sacrifice of the cycling-encoding is its low efficiency: multiple pixels will be used to represent one gray scale level.

Obviously, incoherent correlators are immune to coherent noise. However, it is difficult for incoherent correlators to achieve Fourier transform and subtract operation. Furthermore, some electro-optical device such as high speed FLC SLM works only with polarized coherent beam. Such a device requirement may limit the implementation of incoherent optical correlators.

It is interesting to note that in most optical information processing system the source data to be processed usually comes from a computer. A spatial light modulator (SLM) is then used to convert the data from the electronic to the optical domain. The bottleneck of the processing speed of such an optical system comes from the low data transfer rate from computer memory to SLM and/or the low frame rate of the SLM.

Due to the spectrum response limitation of holographic recording materials, volume holographic correlators usually require coherent beams with a wavelength in the blue-green range. So far lasers operating in this spectrum range and having long coherent length for the implementation of holography are relatively expensive and bulky.

5.4 References for other optical correlators

1. A. Vander Lugt, "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory, IT-10, pp.139-145, 1964.
2. D. Casasent, D. Psaltis, "Position, rotation and scale invariant optical correlation," Appl. Opt., Vol. 15, pp.1795-1799, 1976.

3. J. L. Horner, P. D. Gianino, "Phase-only matched filtering," *Appl. Opt.*, Vol. 23, pp.812-816, 1984.
4. L. Cai, Y. Jin, S. Zhou, P. Yeh, N. Marzwell, H. K. Liu, "Translation sensitivity adjustable compact optical correlator and its application for fingerprint recognition," *Opt. Eng.*, Vol. 35, pp.415-422, 1996.
5. C. S. Weaver, J. W. Goodman, "A technique for optically convolving two functions," *Appl. Opt.*, Vol.5, pp.1248-1248, 1966.
6. F. T. S. Yu, X. J. Xu, "A real-time programmable joint transform correlator," *Opt. Commun.*, Vol.52, pp.10-16, 1984.
7. B. Javidi, C. J. Kuo, "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Appl. Opt.* Vol.27, pp.663-665, 1988.
8. L. Z. Cai, Y. R. Wang, H. Wang, "High efficiency hybrid joint transform correlator with preprocessed input image and replicated binary spectrum array," *J. Mod. Opt.*, Vol.44, pp.1197-1204, 1997.
9. J. Fu, M. P. Schamschula, H. J. Caulfield, "Optical parallel database management system for page oriented holographic memories," *Optics Express*, Vol.5, No.12, p.273-285, 1999.
10. Burr, G.W.; Kobras, S.; Hanssen, H.; Coufal, H., "Content-addressable data storage by use of volume holograms," *Applied Optics*, Vol.38, No.32, p.6779-6784, 1999.
11. P. A. Mitkas,, G. A. Betzos, S. Mailis and N. A. Vainos, "Characterization of associative recall in a volume holographic database system for multimedia applications," *SPIE*, Vol.3388, pp.198-208, 1998.
12. H. K. Liu, Y. H. Jin, N. I. Marzwell, "Advanced ultra-high-capacity optical random access memory and pattern recognition techniques," *Opt. Eng.*, Vol.37, pp.779-788, 1988.
13. Luzhong Cai, Hua-Kuang Liu, "Massive parallel optical pattern recognition and retrieval via a two-stage high-capacity multichannel holographic random access memory system," *Opt., Eng.*, Vol.39, No.2, pp.401-408, 2000.
14. Cheng, G.; Jin, G.; Wu, M.; Liu, H.; He, Q.; Yuan, S., "Compact hybrid optoelectrical unit for image processing and recognition," *SPIE*, Vol.3389, pp.113-121, 1998.
15. H. Liu, M. Wu, G. Jin, Q. He and Y. Yan, "Improving the processing speed of an optoelectronic morphological industrial inspection processor," *SPIE*, Vol.3652, pp.216-223, 1999.

6. Content-addressable processor

Conventional database and volume holographic correlator not only can do associative search, they also store all source data. Different from these two bi-functional systems, content-addressable processor performs only content comparison or other logic operation (such magnitude-comparison as great than or less than) on the input data. It does not store any information. The data to be processed comes from either a control computer or a secondary storage device. That is why we give it the name "processor". By the way, the joint transform correlator and shadow-casting incoherent optical correlator mentioned above do not have any storage capability either. Here we present two types of content-addressable processor.

6.1 Totally digital and electronic processor^[1]

This kind of processor is in fact an integrated semiconductor chip. It has been used as address filters or address translators in ATM and Ethernet based systems. Advanced processors developed in recent years have the parallel search ability. This means the comparison between the input data and all stored address data can be completed in one cycle. NetLogic Microsystems Inc, Mountain View, CA, for example, is providing such kind of commercial products. Using NetLogic's chip, for a 16K×64 routing table for a 12 port Ethernet system, the address filtering and source and destination search time could be less than 30ns. Because of the circuit complexity and limited storage density of this kind of content-addressable processor, usually the capacity is very small. Detailed information on NetLogic's product information can be found at <http://www.netlogicmicro.com/>.

6.2 Optical processor^[2-6]

A good example of an optical content-addressable processor is the one recently proposed by A. Louri et al at ECE Dept, the University of Arizona, which is called multiwavelength optical content-addressable parallel processor (MW-OCAPP). The MW-OCAPP is designed to provide efficient parallel data retrieval and processing by means of moving the bulk of database operations from electronics to optics. Polarization and wavelength-encoding have been proposed to enhance processing parallelism. 11 relational database primitive operations can be realized in MW-OCAPP. The problems faced by this kind of optical processor are 1. the difficulty in manufacturing high speed SLMs, 2. the low data transfer rate from computer to SLM, 3. the computation inaccuracy and 4. bulk system size.

6.3 References for content-addressable processor

1. Ron Neale, "Is content addressable memory (CAM) the key to network success?" Electronic engineering, pp.9-12, February 1999.
2. P. Y. Choo, A. Detofsky, A. Louri, "Multiwavelength optical content-addressable parallel processor for high-speed parallel relational database processing," Applied Optics, Vol.38, No.26, pp.5594-5604, 1999.

3. Ahmed Louri, "Optical content-addressable parallel processor: architecture, algorithms, and design concepts," *Appl. Opt.*, Vol.31, No.17, pp.3241-3258, 1992.
4. Abram Detofsky, Peng Yin Choo, Ahmed Louri, "Equivalency processing parallel photonic integrated circuit (EP³IC): equivalence search module based on multiwavelength guided-wave technology," *Appl. Opt.*, Vol.38, No.32, pp.1-9, 1999.
5. Abram Detofsky, Peng Yin Choo, Ahmed Louri, "Optical implementation of a constant-time multicomparand bit-parallel magnitude-comparison algorithm using wavelength- and polarization-division multiplexing with application to parallel database processing," *Opt. Lett.*, Vol.23, No.17, pp.1372-1374, 1998.
6. Ahmed Louri, James A. Hatch Jr., "An optical associative parallel processor for high-speed database processing: theoretical concepts and experimental results," *Computer*, pp.65-72, November 1994.

**MISSION
OF
AFRL/INFORMATION DIRECTORATE (IF)**

*The advancement and application of Information Systems Science
and Technology to meet Air Force unique requirements for
Information Dominance and its transition to aerospace systems to
meet Air Force needs.*